CLAIMS

What is claimed is:

1. A method, comprising:

    receiving a text sentence comprising a plurality of words, each of the plurality of

    words having a part of speech (POS) tag;

    generating a POS sequence based on the POS tag of each of the plurality of words;

    detecting a prosodic phrase break through a recurrent neural network (RNN), based on

    the POS sequence; and

    generating a prosodic phrases boundary based on the prosodic phrase break.

2. The method of claim 1, further comprising:

    assigning a POS tag for each of the plurality of words of the sentence; and

    classifying the POS tag for each of the plurality of words to a predetermined class.

3. The method of claim 2, wherein the classification of the POS tag comprises adjective,

    adverb, noun, verb, and number.

4. The method of claim 3, wherein the classification of the POS tag further comprises

    quantifier, preposition, conjunction, idiom, and punctuation.

5. The method of claim 1, further comprising segmenting the sentence into the plurality of

    words.

6.   The method of claim 1, wherein detecting the prosodic phrase break through the RNN

      network comprises:

           initializing the RNN network;

           retrieving a POS tag from the tag sequence;

           inputting the POS tag to the RNN network;

           generating an output phrase break associated with the POS tag, from the RNN

                network;

           retrieving a next POS tag from the tag sequence; and

           repeating inputting the POS tag, generating an output phrase break, and retrieving a

                next POS tag, until there are no more POS tags to be processed in the tag

                sequence.

7.   The method of claim 6, further comprising:

           initializing and inputting a first initial phrase break to a first input of the RNN

                network;

           initializing and inputting a first initial POS tag to a second input of the RNN network;

           initializing and inputting a second initial phrase break to a third input of the RNN

                network;

           inputting the first POS tag of the tag sequence to a fourth input of the RNN network;

                and

           inputting the second POS tag of the tag sequence to a fifth input of the RNN network.

8.   The method of claim 7, further comprising:

           inputting the second initial phrase break to the first input of the RNN network;

           inputting the first POS tag of the tag sequence to the second input of the RNN

                network;

inputting the output phrase break, previously generated through the RNN network, to the third input of the RNN network;

inputting the second POS tag of the tag sequence to the fourth input of the RNN network;

inputting the next POS tag from the tag sequence to the fifth input of the RNN network; and

generating a next phrase break associated with the next POS tag through the RNN network.

9. The method of claim 1, wherein the phrase break is generated based on the previously inputted POS tags and previously generated phrase breaks, through the RNN network.

10. A method, comprising:

initializing the RNN network;

retrieving a POS tag from the tag sequence;

inputting the POS tag to the RNN network;

generating an output phrase break associated with the POS tag, from the RNN network;

retrieving a next POS tag from the tag sequence; and

repeating inputting the POS tag, generating an output phrase break, and retrieving a next POS tag, until there are no more POS tags to be processed in the tag sequence.

11. The method of claim 10, further comprising:

initializing and inputting a first initial phrase break to a first input of the RNN network;

initializing and inputting a first initial POS tag to a second input of the RNN network;

initializing and inputting a second initial phrase break to a third input of the RNN network;

inputting the first POS tag of the tag sequence to a fourth input of the RNN network; and

inputting the second POS tag of the tag sequence to a fifth input of the RNN network.

12.   The method of claim 11, further comprising:

inputting the second initial phrase break to the first input of the RNN network;

inputting the first POS tag of the tag sequence to the second input of the RNN network;

inputting the output phrase break, previously generated through the RNN network, to the third input of the RNN network;

inputting the second POS tag of the tag sequence to the fourth input of the RNN network;

inputting the next POS tag from the tag sequence to the fifth input of the RNN network; and

generating a next phrase break associated with the next POS tag through the RNN network.

13.   The method of claim 10, wherein the phrase break is generated based on the previously inputted POS tags and previously generated phrase breaks, through the RNN network.

14.   An apparatus, comprising:

an interface to receive a text sentence comprising a plurality of words, each of the plurality of words having a part of speech (POS) tag;

a text processing unit to generate a POS sequence based on the POS tag of each of the plurality of words;

an recurrent neural network (RNN) to detect a prosodic phrase break, based on the POS sequence and generating a prosodic phrases boundary based on the prosodic phrase break; and

a speech processing unit to perform speech analysis on the prosodic phrase breaks and generating an output speech based on the prosodic phrase breaks.

15. The apparatus of claim 14, wherein the text processing unit assigns the POS tag for each of the plurality of words of the sentence, and classifies the POS tag for each of the plurality of words to a predetermined class.

16. The apparatus of claim 14, wherein the RNN network comprises:

an input layer for receiving input data, the input layer comprising:

a first input to receive a first initial phrase break;

a second input to receive a first initial POS tag;

a third input to receive a second initial phrase break;

a fourth input to receive a first POS tag of the tag sequence; and

a fifth input to receive a second POS tag of the tag sequence;

a hidden layer to perform a prosodic phrase break detection; and

an output layer to generate a prosodic phrase break.

17. The apparatus of claim 16, wherein:

the first input receives the second initial phrase break;

the second input receives the first POS tag of the tag sequence;

the third input receives the output phrase break, previously generated;

the fourth input receives the second POS tag of the tag sequence; and

the fifth input receives the next POS tag from the tag sequence.

18. The apparatus of claim 14, wherein the phrase break is generated based on the previously inputted POS tags and previously generated phrase breaks, through the RNN network.

19. A machine-readable medium having stored thereon executable code which causes a machine to perform a method, the method comprising:

    receiving a text sentence comprising a plurality of words, each of the plurality of words having a part of speech (POS) tag;

    generating a POS sequence based on the POS tag of each of the plurality of words;

    detecting a prosodic phrase break through a recurrent neural network (RNN), based on the POS sequence; and

    generating a prosodic phrases boundary based on the prosodic phrase break.

20. The machine-readable medium of claim 19, wherein the method further comprises:

    assigning a POS tag for each of the plurality of words of the sentence; and

    classifying the POS tag for each of the plurality of words to a predetermined class.

21. The machine-readable medium of claim 19, wherein detecting the prosodic phrase break through the RNN network comprises:

    initializing the RNN network;

    retrieving a POS tag from the tag sequence;

    inputting the POS tag to the RNN network;

generating an output phrase break associated with the POS tag, from the RNN
network;

retrieving a next POS tag from the tag sequence; and

repeating inputting the POS tag, generating an output phrase break, and retrieving a
next POS tag, until there are no more POS tags to be processed in the tag
sequence.

22. The machine-readable medium of claim 21, wherein the method further comprises:

initializing and inputting a first initial phrase break to a first input of the RNN
network;

initializing and inputting a first initial POS tag to a second input of the RNN network;

initializing and inputting a second initial phrase break to a third input of the RNN
network;

inputting the first POS tag of the tag sequence to a fourth input of the RNN network;
and

inputting the second POS tag of the tag sequence to a fifth input of the RNN network.

23. The machine-readable medium of claim 22, wherein the method further comprises:

inputting the second initial phrase break to the first input of the RNN network;

inputting the first POS tag of the tag sequence to the second input of the RNN
network;

inputting the output phrase break, previously generated through the RNN network, to
the third input of the RNN network;

inputting the second POS tag of the tag sequence to the fourth input of the RNN
network;

inputting the next POS tag from the tag sequence to the fifth input of the RNN

network; and

generating a next phrase break associated with the next POS tag through the RNN

network.

24. A machine-readable medium having stored thereon executable code which causes a

machine to perform a method, the method comprising:

initializing the RNN network;

retrieving a POS tag from the tag sequence;

inputting the POS tag to the RNN network;

generating an output phrase break associated with the POS tag, from the RNN

network;

retrieving a next POS tag from the tag sequence; and

repeating inputting the POS tag, generating an output phrase break, and retrieving a

next POS tag, until there are no more POS tags to be processed in the tag

sequence.

25. The machine-readable medium of claim 24, wherein the method further comprises:

initializing and inputting a first initial phrase break to a first input of the RNN

network;

initializing and inputting a first initial POS tag to a second input of the RNN network;

initializing and inputting a second initial phrase break to a third input of the RNN

network;

inputting the first POS tag of the tag sequence to a fourth input of the RNN network;

and

inputting the second POS tag of the tag sequence to a fifth input of the RNN network.

26. The machine-readable medium of claim 25, wherein the method further comprising:

inputting the second initial phrase break to the first input of the RNN network;

inputting the first POS tag of the tag sequence to the second input of the RNN network;

inputting the output phrase break, previously generated through the RNN network, to the third input of the RNN network;

inputting the second POS tag of the tag sequence to the fourth input of the RNN network;

inputting the next POS tag from the tag sequence to the fifth input of the RNN network; and

generating a next phrase break associated with the next POS tag through the RNN network.

27. The machine-readable medium of claim 24, wherein the phrase break is generated based on the previously inputted POS tags and previously generated phrase breaks, through the RNN network.